DOCUMENT RESUME

ED 465 767 TM 033 906

AUTHOR	Dimitrov, Dimiter M.
TITLE	Error Variance of Rasch Measurement with Logistic Ability Distributions.
PUB DATE	2002-04-00
NOTE	25p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Ability; *Error of Measurement; *Item Response Theory; *Test Items; True Scores
IDENTIFIERS	Approximation (Statistics); Calibration; *Rasch Model; *Variance (Statistical)

ABSTRACT

Exact formulas for classical error variance are provided for Rasch measurement with logistic distributions. An approximation formula with the normal ability distribution is also provided. With the proposed formulas, the additive contribution of individual items to the population error variance can be determined without knowledge of the other test items. This feature, not available with previous treatments of classical error variance, may have useful applications in test analysis and development. Formulas for the population true score of individual items are also provided for logistic and normal ability distributions. Thus, from a bank of Rasch calibrated items, one can select items to develop a test with a prespecified: (1) standard error of measurement; (2) mean true score; or (3) error-to-true score ratio for a target population of examinees. These parameters can be used also in comparing subsets of test items that are grouped by substantive or measurement characteristics (e.g., content areas or strands of learning outcomes in proficiency testing). (Contains 4 figures, 4 tables, and 19 references.) (Author/SLD)

Error Variance of Rasch Measurement with Logistic Ability Distributions

Dimiter M. Dimitrov
Kent State University

Dimiter Dimitrov
507 White Hall
College of Education
Kent State University
Kent, OH 44242-0001

Ph. (330) 672-0582
Fax: 9330) 672-3737
E-mail: ddimitro@kent.edu

Paper Presented at the Annual Meeting of the American Educational Research Association April 1-5, 2002, New Orleans, Louisiana

Abstract

Exact formulas for classical error variance are provided for Rasch measurement with logistic distributions. An approximation formula with the normal ability distribution is also provided. With the proposed formulas, the additive contribution of individual items to the population error variance can be determined without knowledge of the other test items. This feature, not available with previous treatments of classical error variance, may have useful applications in test analysis and development. Formulas for the population true score of individual items are also provided for logistic and normal ability distributions. Thus, from a bank of Rasch calibrated items, one can select items to develop a test with a prespecified (a) standard error of measurement, (b) mean true score, or (c) error-to-true score ratio for a target population of examinees. These parameters can be used also in comparing subsets of test items that are grouped by substantive or measurement characteristics (e.g., content areas or strands of learning outcomes in proficiency testing).

Error Variance of Rasch Measurement with Logistic Ability Distributions

In item response theory (IRT), the accuracy of measurement with a test varies across the levels of a latent trait (*ability*), θ, that determines the chances for success on any item in the test. The IRT error variance at a specific θ, $\mathrm{Var}(\hat{\theta}|\theta)$, is inversely related to the information provided by the test at θ (Birnbaum, 1968). This error variance indicates the precision with which ability is estimated at the θ level and is not to be confused with the raw-score variance at θ, $\mathrm{Var}(x|\theta)$. In fact, averaging $\mathrm{Var}(x|\theta)$ over the θ values for a population of examinees yields the classical error variance for this population, σ_e^2, whereas averaging $\mathrm{Var}(\hat{\theta}|\theta)$ results in the so-called *marginal error variance*, $\overline{\sigma}_{e^*}^2$ (Green, Bock, Humphreys, Linn, & Rechase, 1984; Thissen, 1990). While the classical error variance relates to the classical reliability, $\rho_{xx} = 1 - \sigma_e^2 / \sigma_x^2$, the marginal error variance relates to a *marginal reliability*: $\overline{\rho} = 1 - \overline{\sigma}_{e^*}^2 / \sigma_\theta^2$ (Thissen, 1990, p. 167).

In computerized adaptive testing (CAT), for example, the marginal error variance is used with the *stopping rule* according to which "all examinees are tested to the same value of the error variance over as wide a range of ability as practical" (Green et al., 1984, p. 352). Thissen (1990) noted that "marginal reliabilities provide the only comparison between the internal-consistency reliability of a CAT and previously or alternatively used paper-and-pencil forms, for which only classical reliability estimates are available" (p. 167). Such comparisons require quality estimates of the classical error variance for a population of examinees. Moreover, classical error variance and reliability estimates are still widely used in substantive and measurement studies even when IRT information is available (e.g., with standardized tests).

Under classical test theory, the reliability is defined as the ratio of true score variance to observed score variance ($\rho_{xx} = \sigma_\tau^2 / \sigma_x^2$). Although the reliability coefficient is a convenient unitless number between 0 and 1, the classical error variance and its square root (standard error of measurement, SEM) relate to the meaning of the scale and are, therefore, more useful for score interpretations (e.g., Feldt & Brennan, 1989; Thissen, 1990). However, there are methodological and accuracy problems with the classical formula for error variance, $\sigma_e^2 = \sigma_x^2(1 - \rho_{xx})$. In practical applications of this formula, σ_x^2 is replaced with its sample estimate for the data at hand and ρ_{xx} is replaced with the Cronbach's alpha (Cronbach, 1951) or other not sufficiently accurate coefficients. As a reminder, the Cronbach's alpha may underestimate ρ_{xx} (when the components of the test are not at least essentially tau-equivalent; Novick & Lewis, 1967) or overestimate ρ_{xx} (when there are correlated errors; Komaroff, 1997). Moreover, while the definition of reliability, ρ_{xx}, requires (explicitly or implicitly) information about the error variance, the error variance (as is shown later in this paper) can be defined and calculated without information about σ_x^2 or ρ_{xx}.

The purpose of this paper is to propose formulas for independent additive contributions of individual items to classical error variance and true scores using item difficulty estimates with the dichotomous Rasch measurement model (RM) (Rasch, 1960). It should be noted that the approach in this paper uses RM information to deal with accuracy of measurement in the original (number-right score) scale and is not to be confused with Rasch measurement methods that deal with accuracy of ability scores on the logit scale (e.g., Smith, Jr., 2001; Wright & Stone, 1979).

Theoretical Framework

For dichotomously scored items, Lord (1980, p. 52) presented the error variance for a test of *n* items as the mean of the conditional error variances for the number-right scores at the ability estimates of *N* examinees, $\theta_1, ..., \theta_N$:

$$\hat{\sigma}_e^2 = \frac{1}{N}\sum_{j=1}^{N}\sigma_{x|\theta_j}^2 = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n}P_i(\theta_j)[1 - P_i(\theta_j)], \tag{1}$$

where $P_i(\theta_j)$ is the probability for correct answer on item *i* from an examinee with ability θ_j and the product $P_i(\theta_j)[1 - P_i(\theta_j)]$ is the conditional error variance for item *i* at θ_j. The error variance in (1) is based on discrete ability values, $\theta_1, ..., \theta_N$, for the sample. The population error variance is obtained by replacing the summation in Equation 1 for integration over the ability interval:

$$\sigma_e^2 = \sum_{i=1}^{n}\int_{-\infty}^{\infty}P_i(\theta)[1 - P_i(\theta)]\varphi(\theta)d\theta, \tag{2}$$

where $\varphi(\theta)$ is the probability density function (*pdf*) for the ability. The additive contribution of the *i*th test item to the classical error variance in Equation 2 is referred to here as *error variance component*, $\sigma^2_{e_i}$, for this item:

$$\sigma^2_{e_i} = \int_{-\infty}^{\infty} P_i(\theta)[1 - P_i(\theta)]d\theta \tag{3}$$

It is important to note that the additive contribution of individual items to the classical error variance, $\sigma^2_{e_i}$, can be determined without knowledge of the other test items. This feature, not available with traditional treatments of classical error variance, can be practically useful in test development and analysis. For example, given set of Rasch calibrated items, one can select items to develop a test with prespecified standard error of measurement (SEM) that relates the test to specific ability distributions (e.g., logistic, normal, or triangular).

Previous research provides very limited (approximation-based) applications of Equation 2 mostly because of technical difficulties with the integral evaluations. Such evaluations can be performed, for example, by Gaussian quadrature (Bock & Lieberman, 1970), but exact formulas are either difficult or not possible to derive with the $\varphi(\theta)$ for most practically occurring ability distributions. May & Nicewander (1993) used Equation 2, approximating compound binomial conditional distributions of raw scores, with the purpose to compare reliability for number-right scores and percentile ranks with the 3-parameter logistic model for normal and triangular ability distributions.

This paper provides formulas for independent additive contributions of individual items to the classical error variance and true scores as a function of item difficulty estimates with the RM for logistic and normal ability distributions. With the dichotomous RM,

$$P_i(\theta) = \frac{e^{\theta - b_i}}{1 + e^{\theta - b_i}}, \tag{4}$$

where b_i is the difficulty parameter of item *i*. Also, the product $P_i(\theta)[1 - P_i(\theta)]$ equals the first derivative of $P_i(\theta)$:

$$P_i(\theta)[1 - P_i(\theta)] = \frac{e^{\theta - b_i}}{(1 + e^{\theta - b_i})^2} \tag{5}$$

Thus, the error variance component in Equation 3 can be written as

$$\sigma^2_{e_i} = \int_\alpha^\beta [\partial P_i(\theta)/\partial\theta]\varphi(\theta)d\theta = \int_\alpha^\beta \varphi(\theta)dP_i(\theta) \tag{6}$$

The next sections provide exact formulas for the classical error variance components, $\sigma^2_{e_i}$, when $P_i(\theta)$ is with the dichotomous RM and $\varphi(\theta)$ is the *pdf* for logistic and normal distributions.

Error Variance Components for Logistic Ability Distribution

The *pdf* of a logistic distribution (e.g., Evans, Hastings, & Peacock, 1993, p. 98) with the location at the origin of the scale is

$$\varphi(\theta) = \frac{\exp(\theta / c)}{c[1 + \exp(\theta / c)]^2} \tag{7}$$

where *c* is the scale parameter. Figure 1 shows the *pdf* of logistic distributions with $c = 1/2$ and 1. These two specific logistic shapes were selected because (a) their logistic *pdf*s lead to exact integral evaluations for error variance components and (b) the shapes that they produce capture normal-like ability distributions that may practically occur with Rasch measurement.

As one may also notice, $\varphi(\theta)$ in Equation 7 is the first derivative of the function

$$\Phi(\theta) = \frac{\exp(\theta / c)}{1 + \exp(\theta / c)}$$

Thus, replacing $\varphi(\theta)$ in Equation 5 with the first derivative of $\Phi(\theta)$, we have

$$\sigma^2_{e_i} = \int_{-\infty}^{\infty} [\partial P_i(\theta)/\partial\theta][\partial\Phi(\theta)/\partial\theta]d\theta = \int_{-\infty}^{\infty} [\partial P_i(\theta)/\partial\theta]d\Phi(\theta)$$

Then, with integration by parts and simple calculations, we have

$$\sigma^2_{e_i} = [\partial P_i(\theta)/\partial\theta]\Phi(\theta)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\Phi(\theta)d[\partial P_i(\theta)/\partial\theta]$$

$$= 0 - \int_{-\infty}^{\infty}\Phi(\theta)[\partial^2 P_i(\theta)/\partial\theta^2]d\theta$$

$$= -\int_{-\infty}^{\infty}\frac{\exp(\theta/c)\exp(\theta-b_i)[1-\exp(\theta-b_i)]}{[1+\exp(\theta/c)][1+\exp(\theta-b_i)]^3}d\theta$$

Let us denote $E_i = e^{b_i}$. Using the substitution rule for integration with $x = \exp(\theta)$, we have

$$\sigma^2_{e_i} = \int_0^{\infty}\frac{E_i x^{1/c}(x-E_i)}{(1+x^{1/c})(x+E_i)^3}dx \tag{8}$$

The evaluations of the integral in Equation 8 for $c = 1$ and $c = 1/2$, respectively, are:

1. With $c = 1$,

$$\sigma^2_{e_i} = \frac{E_i(b_iE_i - 2E_i + b_i + 2)}{(E_i - 1)^3} \tag{9}$$

When $b_i = 0$, the denominator of the ratio in Formula 9 equals zero. In this case, calculating the the limit of the ratio at $b_i = 0$, we have: $\sigma^2_{e_i} = 0.1667$ (see Table 1).

2. With $c = 1/2$,

$$\sigma^2_{e_i} = \frac{E_i\left[8(1-b_i)E_i^3 + 8(b_i+1)E_i + \pi E_i^4 - 6\pi E_i^2 + \pi\right]}{2(E_i^2+1)^3} \tag{10}$$

The integral evaluations in Formulas 9 and 10 were verified using MATLAB 5.3 (MathWorks, Inc., 1999). Thus, when the ability scores follow a logistic distribution located at the origin of the scale with a scale parameter $c = 1$ or 1/2, the error variance components for individual items can be calculated with Formulas 9 or 10, respectively. Values of error variance components, $\sigma^2_{e_i}$, are tabulated in Table 1 and graphed in Figure 2. Since $\sigma^2_{e_i}(-b_i) = \sigma^2_{e_i}(b_i)$, one can use Table 1 for any set of Rasch calibrated items, with b_i ranging from -8 to 8. For example, the error variance component at $b_i = 1$ for the logistic *pdf* with $c = 1$ (0.1509) equals the error variance component at $b_i = -1$ for the same *pdf.*

Error Variance Components for Normal Ability Distribution

With $\varphi(\theta)$ for the standard normal *pdf*, the integral for the error variance component in Equation 6 can be written as

$$\sigma^2_{e_i} = \int_{-\infty}^{\infty} \frac{\exp(\theta - b_i)}{[1 + \exp(\theta - b_i)]^2} \left(\frac{1}{\sqrt{2\pi}} \exp(-.5\theta^2 \right) d\theta \qquad (11)$$

Since a closed form evaluation of the integral in Equation 11 does not exist, an approximation formula for was developed in two steps. First, using the computer program MATLAB 5.3 (MathWorks, Inc., 1999), quadrature method evaluations of the integrals for $\sigma^2_{e_i}$ were obtained. The results are tabulated in Table 2 and graphed in Figure 3. Second, it was found that the values of $\sigma^2_{e_i}$ fit the following approximation:

$$\hat{\sigma}^2_{e_i} = A + B\exp[-0.5(b_i / C)^2], \qquad (12)$$

where:

(a) $A = 0.011$, $B = 0.195$, and $C = 1.797$, if $|b_i| < 4$, or

(b) $A = 0.0023$, $B = 0.171$, and $C = 2.023$, if $|b_i| \geq 4$.

Note that $\sigma^2_{e_i}$ is an even function of the Rasch item difficulty, i.e., $\sigma^2_{e_i}(-b_i) = \sigma^2_{e_i}(b_i)$. Thus, one can use Table 2 for items with Rasch difficulty that range in from -6 to 6 on the logit scale. The absolute error of approximation with Formula 12, $|\varepsilon|$, ranges from 0 to 0.0008, with a mean of 0.00021 and a standard deviation of 0.00017. Also, the approximation errors, ε, practically do not affect the total classical error variance because they vary in sign (see Table 2) and cancel out to a large degree in summing error variance components. One can use Formula 12 with any normal ability distribution, $N(\mu_\theta; \sigma_\theta)$, after appropriate transformation of the Rasch item difficulties for the test: $b_i^* = (b_i - \mu_\theta)/\sigma_\theta$ (e.g., Lord, 1980).

Item True Score as a Function of Rasch Item Difficulty

In this section, the expected true score for individual items, τ_i, is represented as a function of their Rasch difficulty with the normal and logistic ability distributions (c = 1 and 1/2):

$$\tau_i = \int_{-\infty}^{\infty} \tau_{i|\theta}\varphi(\theta)d\theta = \int_{-\infty}^{\infty} P_i(\theta)\varphi(\theta)d\theta, \qquad (13)$$

where $P_i(\theta)$ is with the dichotomous Rasch measurement model and $\varphi(\theta)$ is the ability *pdf*. Then, the total true score for a test of n items is: $\tau = \sum_1^n \tau_i$; (the average true score is $\bar{\tau} = \tau / n$).

Logistic Ability Distribution (scale parameter $c = 1$)

In this case, $\varphi(\theta)$ in Equation 13 is substituted for its expression (with $c = 1$) in Equation 7:

$$\tau_i = \int_{-\infty}^{\infty} \frac{\exp(\theta - b_i)\exp(\theta)}{[1+\exp(\theta - b_i)][1+\exp(\theta)]^2} d\theta \qquad (14)$$

Using the integral substitution $t = \exp(\theta)$, the evaluation of the integral in Equation 14 becomes straightforward and leads to the formula for the item true score in this case:

$$\tau_i = \frac{(b_i - 1)\exp(b_i) + 1}{[\exp(b_i) - 1]^2} \qquad (15)$$

Logistic Ability Distribution (scale parameter c = 1/2)

Now, $\varphi(\theta)$ in Equation 13 is substituted for its expression (with $c = 1/2$) in Equation 7:

$$\tau_i = \int_{-\infty}^{\infty} \frac{2\exp(\theta - b_i)\exp(2\theta)}{[1+\exp(\theta - b_i)][1+\exp(2\theta)]^2} d\theta \tag{16}$$

Again, using the substitution $t = \exp(\theta)$, we obtain the following formula:

$$\tau_i = \frac{\pi \exp(b_i)[\exp(2b_i) - 1] - 2(2b_i - 1)\exp(2b_i) + 2}{2[1+\exp(2b_i)]^2} \tag{17}$$

Normal Ability Distribution

When $\varphi(\theta)$ is the pdf of the standard normal distribution, there is no closed form for the integral evaluation in Equation 13. Using the quadrature method of integration with the computer program MATLAB (MathWorks, Inc, 1999), the following approximation formula was developed for the true score of individual items in this case:

$$\tau_i = -0.0114 + \frac{1.0228}{1+\exp(b_i / 1.226)} \tag{18}$$

with an absolute error smaller than 0.002.

It should be noted also that Formulas 17 and 18 produce almost equal true scores for the same value of b_i (with an absolute difference smaller than 0.008). This is graphically represented in Figure 5 where the item true score curves produced by Formulas 17 and 18 overlap. One can use Formula 18 with any normal distribution, $N(\mu_\theta; \sigma_\theta)$, after the appropriate transformation of the item difficulty estimates $b_i^* = (b_i - \mu_\theta) / \sigma_\theta$ (e.g., Lord, 1980).

The proposed formulas for trues scores of individual items can be useful in theoretical and practical (e.g., simulation) studies or developing computer software for Rasch measurement. Also, given a set of n items, one can calculate the ratio $\overline{\sigma}_e / \overline{\tau}$, where $\overline{\sigma}_e = \sigma_e / n$ and $\overline{\tau} = \tau / n$.

This unitless ratio, referred to here as *error-to-true score ratio* (ETR), can be useful in comparing magnitudes of measurement error across different scale units. Thus, one can control the values of σ_e, τ, and ETR for a set of Rasch calibrated items or compare such values for different subsets of items grouped by some substantive or measurement characteristics (e.g, content area or strands of learning outcomes).

Example

This example illustrates the calculation of error variance components and true scores for Rasch calibrated items with the normal ability distribution. Data were collected with the Ohio Off-Grade Proficiency Test-Mathematics (Riverside Publishing, 1995) for 2547 fifth graders from a large urban area in northeastern Ohio. Using RASCAL (Assessment System Corporation, 1995), it was found that the data fit the Rasch model with 20 dichotomous items and the ability scores (in logits) followed the standard normal distribution. The item difficulty estimates, b_i, error variance components (calculated with Formula 12), and true scores (calculated with Formula 18) are provided in Table 3. The standard errors of the estimated item difficulties, b_i, ranged from 0.04 to 0.06. In Table 3, the items are grouped by their content area: Algebra (nine items), Geometry (four items), and Data Analysis/Probability (six items). The sum of all 20 error variance components in Table 3 is the classical error variance for the test: $\sigma_e^2 = 3.675$. Thus, the standard error of measurement for the population with this test: $\text{SEM} = \sqrt{3.675} = 1.917$.

Figure 4 presents the intervals $\tau_\theta \pm$ SEM and $\tau_\theta \pm \sigma_{e|\theta}$, where τ_θ is the true score for the 20-item test at the ability level θ and $\sigma_{e|\theta}$ is the conditional error variance for the test score at θ. Evidently, the SEM bounds are slightly off the conditional ones at either very low or very high ability levels thus providing a good overall estimation of the error associated with the raw scores on the test. Table 4 provides, by content areas, the values of $\overline{\sigma}_e$, $\overline{\tau}$, and their ratio, ETR. As one can see, the students are most successful on the algebra items ($\overline{\tau} = .534$) and least successful on the data analysis/probability items ($\overline{\tau} = .442$). The lowest error of measurement per item is associated with the algebra items (0.146) and the highest, with the geometry items (0.187). The lowest relative error of measurement (0.273) is associated with the algebra items and the highest (0.389), with the data analysis/probability items. For the whole test, the error of measurement is relatively small compared to the true score (ETR = 0.191).

Conclusion

This paper provides formulas for independent additive contributions of items to the classical error variance and true scores with Rasch measurement for logistic and normal ability distributions. The proposed formulas deal with expected values of classical measurement error and true scores for a population with specified ability distribution and do not require information about specific ability scores for a sample of examinees. The method used in this paper eliminates methodological and accuracy problems related to traditional methods based on Cronbach's alpha or other not sufficiently accurate classical coefficients. Also, it uses RM calibration of items to evaluate the accuracy of number-right scores but is not be confused with Rasch measurement methods that deal with accuracy of ability scores on the logit scale (e.g., Smith, Jr., 2001; Wright & Stone, 1979).

The logistic ability distribution used in this study is located in the origin of the scale, with a scale parameter of 1 and 1/2, respectively. These two specific values relate to bell-shaped ability curves that may practically occur in measurement situations and, most importantly, yield exact integral evaluations in the derived formulas for classical error variance (see Formulas 9 and 10). This is not true with just any scale parameter of the logistic distribution.

Formula 12 generates error variance components of individual items with the standard normal ability distribution. It can be used with any normal ability distribution after standardizing the ability scores and then using the same linear transformation for the Rasch item difficulty estimates. The absolute error of approximation with Formula 12 varies from 0 to 0.0008, with a mean of 0.00021 and a standard deviation of 0.00017. Also, the approximation errors practically do not affect the total classical error variance because they vary in (positive/negative) sign (see

Table 2) thus canceling out to a large degree in the summation of error variance components.

With the proposed formulas, given the RM calibration of items, one can evaluate the classical error variance and SEM for a population without further data collection. The additive contribution of each item to the classical error variance can be determined without knowledge of the other items in the test. This feature, not available with previous treatments of classical error variance, can be very useful in test development and analysis. For example, from a bank of items calibrated with the dichotomous RM, one can select items to develop a test with a desirable error variance. One can use the tabulations in Tables 1 and Table 2 or perform calculations with the appropriate formula using basic statistical programs, spreadsheet programs, or even calculators.

It should be noted that skewed ability distributions also occur with Rasch measurement (e.g., in medical studies; Wright, 2000). Dimitrov (2000) provided exact formulas for error variance components with skewed (e.g. triangular and pseudo-lognormal) ability distributions with Rasch measurement.

Formulas 15, 17, and 18 provide the true score of individual items as a function of their Rasch item difficulty for two logistic ability distributions (with scale parameters $c = 1$ and 1/2) and the standard normal ability distribution, respectively. Thus, given a bank of Rasch calibrated items, one can control the error of measurement, true score, and error-to-true score ratio for any set of items. One can also compare measurement errors and trues scores for different subsets of test items that are grouped by substantive or measurement characteristics (e.g., content area and strands of learning outcomes).

In conclusion, using Rasch measurement information for the evaluation of independent and additive contribution of individual items to classical error variance and true scores provides high quality in understanding, calculating, and reporting classical error of measurement.

References

Assessment System Corporation (1995a). *User's Manual for RASCAL Rasch analysis program (Windows version 3.5)*. St. Paul, MN: Author.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick, *Statistical theories of mental test scores* (chapters 17-20). Reading, MA: Addison-Wesley.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for *n* dichotomously scored items. *Psychometrika, 35*, 179-197.

Cronbach, L. J. (1951). Coeficient alpha and the internal structure of a test. *Psychometrika, 16*, 297-334.

Dimitrov, D. M. (2001, October). Reliability of Rasch measurement with skewed ability distributions. Paper presented at the International Conference on Objective Measurement. Chicago, IL.

Evans, M.., Hastings, N., & Peacock, B. (1993). *Statistical distributions* (2nd ed.). NY: John Wiley.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (pp. 105-146). New York: Macmillan.

Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement, 21*, 347-360.-

Komaroff, E. (1997). Effect of simultaneous violations of essential tau-equivalent and uncorrelated errors on coefficient alpha. *Applied Psychological Measurement, 21*, 337-348.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

MathWorks, Inc. (1999). *Learning MATLAB (Version 5.3)*. Natick, MA: Author.

May, K., & Nicewander, W. A. (1993). Reliability and information functions for percentile ranks. *Psychometrika, 58*, 313-325.

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*, 1-13.

Rasch, G. (1960). *Probabilistic models for intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.

Riverside Publishing (1995). *Ohio Off-Grade Proficiency Tests: specifically designed to measure*

Ohio's model course of study. Chicago, IL; Author.

Smith, Jr., E. V. (2001). Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. *Journal of Applied Measurement, 2*, 281-311.

Thissen, D. (1990). Reliability and measurement precision. In H. Winer, *Computerized Adapative Testing. A Primer* (chapter 7). Hillsdale, NJ: Lawrence Erlbaum.

Wright, B. D. (2001). Separation, reliability and skewed distributions. *Rasch Measurement Transactions 14:4*, Chicago: IL.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

Table 1

Error Variance Components ($\sigma^2_{e_i}$) *as a function of Rasch Item Difficulty* (b_i) *for Logistic Ability Distributions with* c = 1 and c = 1/2

b_i	Scale parameter $c = 1$	Scale parameter $c = 1/2$	b_i	Scale parameter $c = 1$	Scale parameter $c = 1/2$
0	.1667	.2146	4.1	.0384	.0226
.1	.1665	.2142	4.2	.0360	.0206
.2	.1660	.2131	4.3	.0337	.0189
.3	.1652	.2113	4.4	.0316	.0172
.4	.1640	.2089	4.5	.0295	.0157
.5	.1626	.2057	4.6	.0276	.0143
.6	.1608	.2019	4.7	.0258	.0131
.7	.1587	.1976	4.8	.0241	.0119
.8	.1564	.1927	4.9	.0225	.0108
.9	.1538	.1874	5.0	.0210	.0099
1.0	.1509	.1816	5.1	.0195	.0090
1.1	.1479	.1755	5.2	.0182	.0082
1.2	.1446	.1690	5.3	.0169	.0074
1.3	.1411	.1624	5.4	.0157	.0067
1.4	.1375	.1555	5.5	.0146	.0061
1.5	.1337	.1486	5.6	.0136	.0056
1.6	.1298	.1415	5.7	.0126	.0050
1.7	.1258	.1345	5.8	.0117	.0046
1.8	.1217	.1275	5.9	.0108	.0042
1.9	.1175	.1205	6.0	.0100	.0038
2.0	.1133	.1137	6.1	.0093	.0034
2.1	.1091	.1070	6.2	.0086	.0031
2.2	.1049	.1005	6.3	.0080	.0028
2.3	.1006	.0942	6.4	.0074	.0026
2.4	.0964	.0881	6.5	.0068	.0023
2.5	.0923	.0823	6.6	.0063	.0021
2.6	.0882	.0767	6.7	.0058	.0019
2.7	.0841	.0713	6.8	.0054	.0017
2.8	.0801	.0662	6.9	.0050	.0016
2.9	.0763	.0614	7.0	.0046	.0014
3.0	.0725	.0569	7.1	.0042	.0013
3.1	.0688	.0526	7.2	.0039	.0012
3.2	.0652	.0486	7.3	.0036	.0010
3.3	.0617	.0448	7.4	.0033	.0010
3.4	.0584	.0413	7.5	.0030	.0009
3.5	.0552	.0380	7.6	.0028	.0008
3.6	.0520	.0349	7.7	.0026	.0007
3.7	.0491	.0320	7.8	.0024	.0006
3.8	.0462	.0294	7.9	.0022	.0006
3.9	.0435	.0269	8.0	.0020	.0005
4.0	.0408	.0247			

Note. For negative b_i, one can use that $\sigma^2_{e_i}(-b_i) = \sigma^2_{e_i}(b_i)$.

Table 2

Error variance components ($\sigma^2_{e_i}$) *and their approximations* ($\hat{\sigma}^2_{e_i}$) *with formula* (13) *as a function of Rasch item difficulty* (b_i).

b_i	$\sigma^2_{e_i}$	$\hat{\sigma}^2_{e_i}$	ε	b_i	$\sigma^2_{e_i}$	$\hat{\sigma}^2_{e_i}$	ε
0.0	.2066	.2060	.0006	3.1	.0554	.0550	.0004
0.1	.2063	.2057	.0006	3.2	.0513	.0509	.0004
0.2	.2054	.2048	.0006	3.3	.0474	.0471	.0003
0.3	.2038	.2033	.0005	3.4	.0437	.0436	.0002
0.4	.2017	.2012	.0005	3.5	.0403	.0403	.0000
0.5	.1990	.1986	.0004	3.6	.0371	.0372	-.0001
0.6	.1957	.1954	.0003	3.7	.0341	.0344	-.0003
0.7	.1920	.1918	.0002	3.8	.0313	.0318	-.0006
0.8	.1877	.1876	.0001	3.9	.0287	.0295	-.0008
0.9	.1830	.1830	.0000	4.0	.0263	.0265	-.0002
1.0	.1779	.1780	-.0001	4.1	.0241	.0242	-.0002
1.1	.1725	.1727	-.0002	4.2	.0220	.0221	-.0001
1.2	.1668	.1670	-.0002	4.3	.0201	.0202	-.0000
1.3	.1608	.1611	-.0003	4.4	.0184	.0184	.0000
1.4	.1546	.1550	-.0003	4.5	.0168	.0167	.0001
1.5	.1483	.1486	-.0004	4.6	.0153	.0152	.0001
1.6	.1418	.1422	-.0004	4.7	.0139	.0138	.0001
1.7	.1353	.1357	-.0004	4.8	.0127	.0125	.0001
1.8	.1288	.1291	-.0003	4.9	.0115	.0114	.0001
1.9	.1222	.1225	-.0003	5.0	.0105	.0104	.0001
2.0	.1158	.1160	-.0002	5.1	.0096	.0094	.0001
2.1	.1094	.1095	-.0001	5.2	.0087	.0086	.0001
2.2	.1031	.1032	-.0001	5.3	.0079	.0078	.0001
2.3	.0970	.0970	.0000	5.4	.0072	.0072	.0000
2.4	.0911	.0909	.0001	5.5	.0065	.0065	-.0000
2.5	.0853	.0851	.0002	5.6	.0059	.0060	-.0001
2.6	.0797	.0795	.0003	5.7	.0054	.0055	-.0002
2.7	.0744	.0741	.0003	5.8	.0049	.0051	-.0002
2.8	.0693	.0689	.0004	5.9	.0044	.0047	-.0003
2.9	.0644	.0640	.0004	6.0	.0040	.0044	-.0004
3.0	.0598	.0594	.0004				

Note. $\varepsilon = \sigma^2_{e_i} - \hat{\sigma}^2_{e_i}$; For negative b_i, use $\sigma^2_{e_i}(-b_i) = \sigma^2_{e_i}(b_i)$.

Table 3

Error Variance Components, $\sigma^2_{e_i}$, *and true scores,* τ_i, *for the Rasch Calibrated Items of the Example Mathematics Test Grouped by Content Areas*

Item	b_i	$\sigma^2_{e_i}$	τ_i
	Algebra		
2	-0.659	0.1933	.634
3	-0.988	0.1786	.696
4	0.600	0.1954	.377
8	-0.648	0.1937	.632
9	-0.101	0.2057	.521
10	1.085	0.1735	.287
15	0.266	0.2039	.445
16	-0.792	0.1880	.660
17	-0.246	0.2042	.551
	Geometry		
1	-1.563	0.1446	.788
6	0.969	0.1796	.309
14	-0.354	0.2023	.573
19	-0.886	0.1837	.677
20	1.241	0.1646	.261
	Data analysis/Probability		
5	-0.275	0.2037	.557
7	-1.119	0.1716	.718
11	1.044	0.1757	.294
12	2.180	0.1044	.136
13	0.315	0.2030	.435
18	-0.069	0.2059	.514

Note. The classical error variance for the test is $\sigma^2_e = \sum_{i=1}^{20} \sigma^2_{e_i} = 3.6754$.

Table 4

Mean SEM per Item ($\overline{\sigma}_e$), True Score ($\overline{\tau}$), and Their Ratio by Content Areas of the Mathematics Test

Content area	n	$\overline{\sigma}_e$	$\overline{\tau}$	$\overline{\sigma}_e / \overline{\tau}$
Algebra	9	0.146	.534	0.273
Geometry	5	0.187	.521	0.359
Data analysis/Probability	6	0.172	.442	0.389
Total test	20	0.096	.503	0.191

Note. n = number of items;

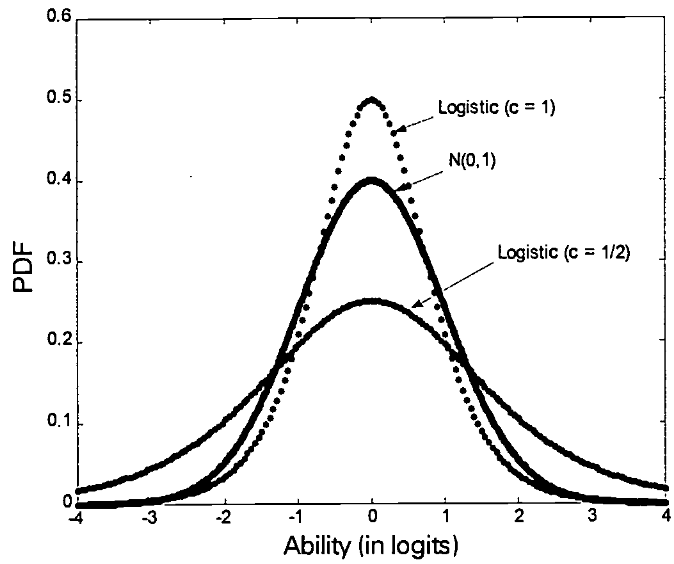

Figure 1. Probability density functions (PDF) of the standard normal distribution and two logistic distributions with scale parameters $c = 1$ and $c = 1/2$.

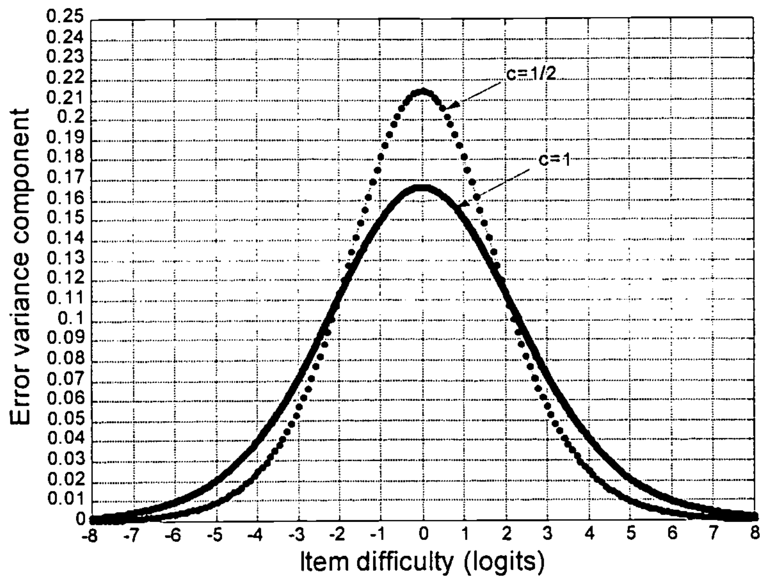

Figure 2. Error variance components, $\sigma^2_{e_i}$, as a function of Rasch item difficulty for logistic ability distributions with scale parameters $c = 1$ and $c = 1/2$.

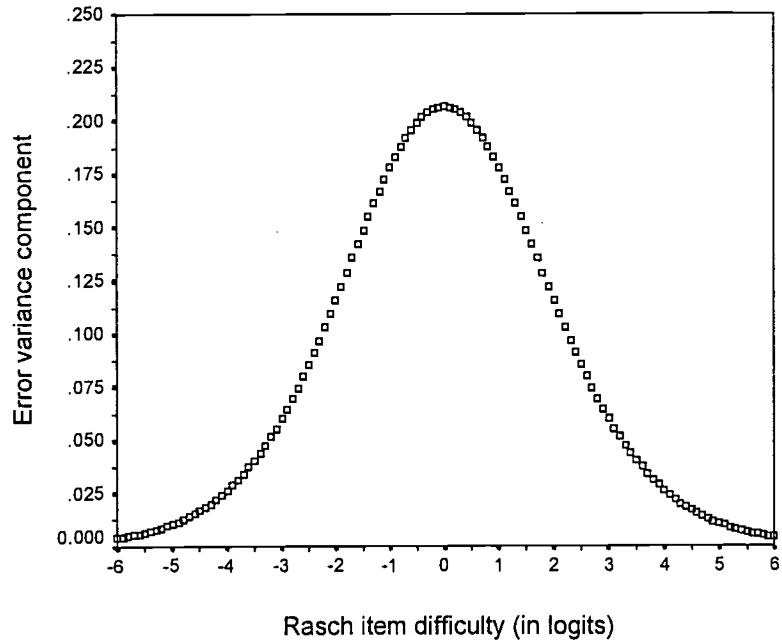

Figure 3. Error variance components, $\sigma^2_{e_i}$, as a function of Rasch item difficulty for the standard normal ability distribution

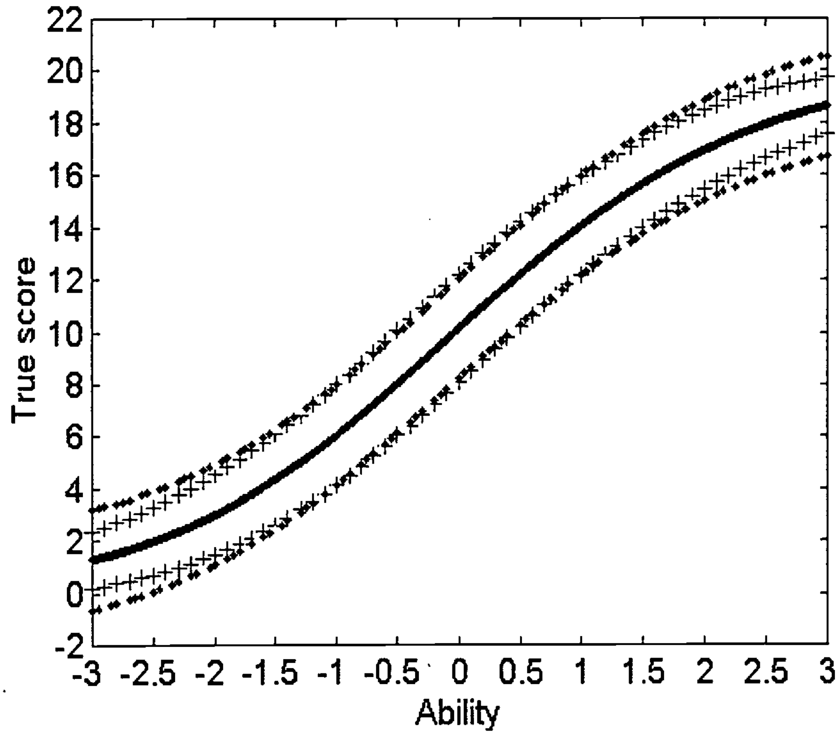

Figure 4. Boundaries for measurement error about conditional true scores, τ_θ, for the 20-item mathematics test: $\tau_\theta \pm$ SEM (dashed lines, --) and $\tau_\theta \pm \sigma_{e|\theta}$ (maker lines, +) ; (in this case, SEM = 1.917).

BEST COPY AVAILABLE

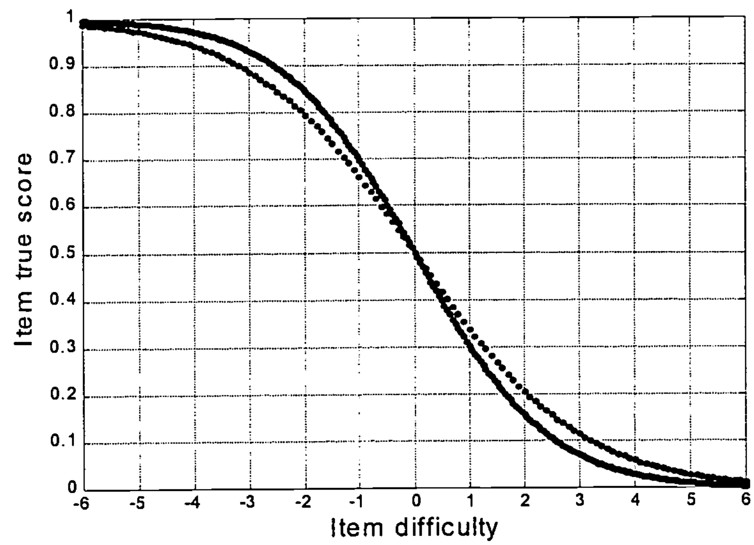

Figure 5. Item true score as a function of Rasch item difficulty when the population ability distribution is normal (solid line), logistic with $c = 1$ (dotted line, and logistic with $c = 1/2$ (dashed line -- superimposed on the solid line)

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM033906

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Error Variance of Rasch Measurement with Logistic Ability Distributions

Author(s): Dimiter M. Dimitrov

Corporate Source: Kent State University

Publication Date: 04/02/02

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: Dimitrov

Printed Name/Position/Title: DIMITER DIMITROV, Ph.D.

Organization/Address: Kent State University, Kent, OH 44242

Telephone: 330-672-0582 FAX: 330 672-3737

E-Mail Address: ddimitro@kent.edu Date: 04/02/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)